International Journal of Scientific & Engineering Research, Volume 7, Issue 5, May-2016
ISSN 2229-5518

Cross resolution face matching based on ensemble co-transfer learning

Renju Rachel Varghese, Dr.Shyama Das, Reeba Mariyam Cherian

Abstract— With the advancement in technology, face recognition has a significant and widespread application in image analysis, identification of persons in law enforcement, identifying the culprit during riots, breach of security, etc. Mainly face recognition algorithms are developed for matching high resolution images, which are being trained in a controlled environment, often fail to match high resolution images with the probe images. In this paper, we are improving cross resolution image matching using the co-transfer learning framework, which uses the technique of co-training and transfer learning in a non-separable manner. Using transfer learning we are transferring the knowledge from the source domain to target domain. The classifiers decision boundary is updated and pseudo labels are being assigned to the unlabeled probe instances by using online co-training of the SVM classifier. The feature vectors are being extracted using local binary pattern and histogram of oriented gradients, and SIFT. Unification of the transfer learning and co-training in the seamless manner helps to raise the operation of the cross resolution face matching.

Index Terms—. Co training, Cross resolution, co-transfer learning, probe images, pseudolabels, source domain, target domain.

1 Introduction

WHILE face recognition research has been exercised for more than three decades and many promising practical face recognition systems have been built up. With the advancements in technology as well as growing installation of surveillance cameras, there is an increasing demand of face recognition technology for surveillance applications, in different areas like enforcement of law and security purposes. Surveillance cameras are mainly designed for maximum coverage from a static location and gives low resolution images. The need to identify the face of the individuals from very low resolution images has emerged as a new covariate in face recognition.

Facial recognition systems can fail, if the optimal conditions are not met. The system can be fooled by hats/beards/sunglasses/face masks. The other factors that can cause impact to the accuracy of matches, includes room lighting, camera angle and difference in resolution (Cross-Resolution).The image capturing probes are of different resolution and this causes the facial recognition systems to result in poor performance. The presence of pose, expression and illumination along with different resolution could further exacerbate the trouble. The focus of the paper is to match low resolution images/probe images with the high resolution images.

In that respect there are various approaches to the problem of cross resolution. They are broadly categorized into two major groups: super resolution and transformation based. Super resolution techniques can be generally classified into 1) Reconstruction based and 2) Learning based. Super resolution technique tries to retrieve the lost high frequency data from low level image primitives[2]. Super resolution methods produce a reconstructed high resolution image from low level image primitives by making an assumption of the image content. Super resolution uses upsampling technique, in order to reconstruct a high resolution image from a low resolution image.

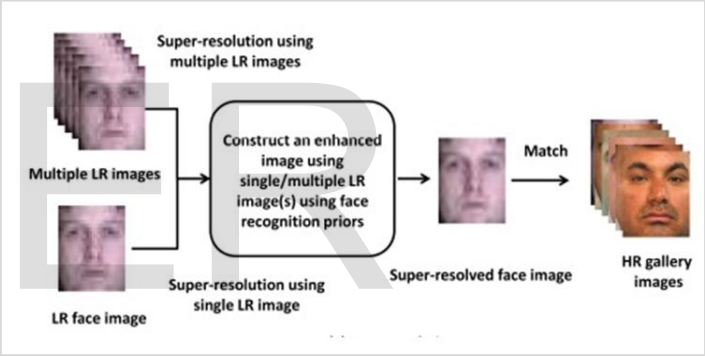


Figure1: Broad view of super resolution face matching

Approach	Technique	Disadvantage	Gallery resolution
Coherent subspace	A coherent subspace of coherent features between the PCA features of HR and LR images.	Poorly cope with nonlinear variations in viewing condition.	72*72/12*12
Multi-modal tensor face	Integrates pixel domain super resolution and recognition by directly computing the maximum likelihood identity.	Due to environmental variations & distortions, failed to significantly improve the recognition problem	56*36/14*9
S2R2	Face features, as they would be extracted for a FR are included in a super-resolution method as prior inform	Fails to cope up with variations in face	24*24/6*6

Table 1: Existing super resolution algorithms

Super resolution based approaches for cross resolution face matching enhances the low quality probe image before recognition[2][3]. Disadvantage of early super resolution techniques are they only considered face images under fixed imaging conditions and don't cope with the nonlinear variations in viewing condition which are away from the training data. The technique of super resolution face matching is shown in Figure 1.Various techniques that are used for face matching is being given in Table 1.

In transformation based approaches, we mainly down sample the images. The disadvantage with this method is, useful information for face recognition such as edges, texture and other frequency information is being compromised while down sampling.[1] The technique of transformation based approaches is shown in figure 2. Various methods that are used for transformation based cross matching is briefly given intable 2.

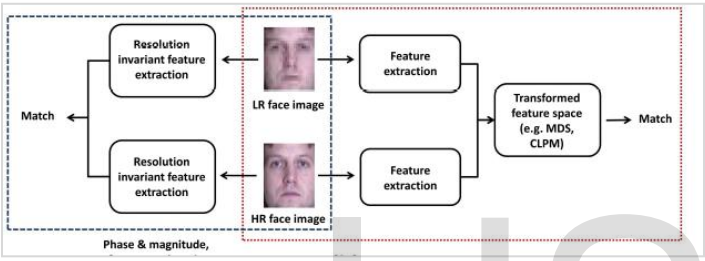


Figure 2: Broad view of Transformation based method for face matching

Approach	Technique	Advantage	Gallery Resolution
LFD	We build the LFD using the local phase quantization & exploring both blur invariant magnitude & phase information in low frequency domain.	The image must be well aligned	64*48/28*24
CLPM	Coupled mappings are learned through optimizing the objective function to minimize the difference between correspondences of LR & HR	Favors task of classification	72*72/12*12
Synthesis based LR face recognition	For each training image use the relighting approach to generate multiple images with different illumination condition. Learn the best dictionaries to represent resolution specific enlarged training matrices	Provides better result	48*40/19*16

Table 2: Existing transformation based algorithms

2. PROBLEM FORMULATION

Another related challenge that pertains to training in the controlled environment, with high resolution images with an ample amount of labeled data with testing in the uncontrolled environment with low resolution images, that has only few labeled data and large amount of unlabeled data. The problem we have in our hands has a large amount of labeled data in the source domain and few labeled data and large amount of unlabeled data in the target domain.

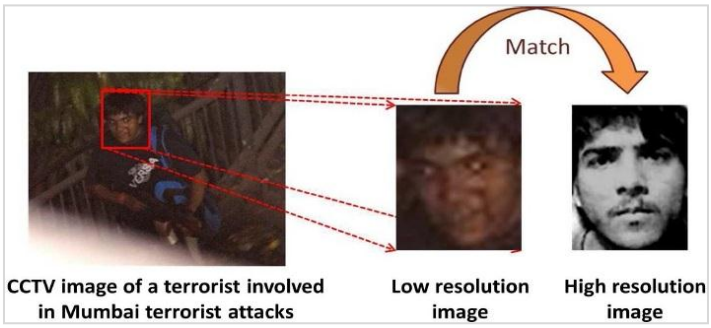


Figure 2: Illustrating the difference in matching a) low resolution and high resolution images.

Under these variations, the performance of existing biometric system degrades, because it is unable to efficiently utilize the knowledge learned in the source domain and there is only few labeled low resolution data which can be used for training the algorithms[1]. There is an abundance of unlabeled low resolution data in the target domain during testing. These unlabeled data must be assigned pseudo labels during the testing phase by using the knowledge learnt from the training phase. In the source domain, high resolution query/probe images are matched with the high resolution gallery whereas in the target domain, low resolution probe /query images are matched with the high resolution gallery. To resolve this problem, we offer a co-transfer learning framework which is the amalgamation of transfer learning and co-training. The knowledge learned in the source domain is transferred for the efficient matching in the target domain using online transfer learning and transfer learning is achieved using co training.

Transfer learning is used to extend the knowledge learned from the source domain to the target domain, to efficiently match the low resolution probe images with the HR gallery in the target domain.

Co-training is used to facilitate transfer learning, by assigning pseudo labels for the unlabeled probe instances from the target domain

3 RELATED WORK: CO-TRANSFER LEARNING FRAMEWORK

The illustration of co-transfer learning framework, which is the amalgamation of transfer learning and co-training is shown in the below figure[1].

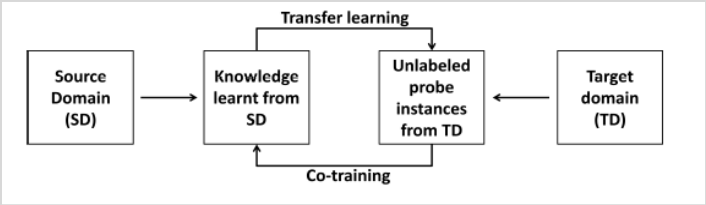


Figure 4: Cross pollination of transfer learning and co-training for transferring knowledge from the source domain to target domain

Transfer learning: Humans can often transfer knowledge learnt previously to novel situations. Transfer learning is motivated with human learning. In machine learning, given some prior knowledge in a related task, traditional algorithms are unable to adapt to a new task in a new domain and have to learn the new task in new domain from the scratch. Transfer learning is the ability of a system to acknowledge and use the knowledge and skills learned in previous tasks to novel tasks (in new domains)[1]. Existing approaches to transfer learning can be categorized as 1) inductive 2) transductive 3) unsupervised transfer learning. Transfer learning can be also categorized in terms of domain representation as homogenous transfer learning and heterogeneous transfer learning. In homogenous transfer learning the source and target domain have the same feature space, whereas in heterogeneous transfer learning the source domain and target domain have different feature space. Most of the transfer learning methods operate in offline mode and assumes that data from the target domain is available up front.

Consider a scenario, where the labeled data in the object domain is limited and obtaining pseudo labels for the target domain is expensive and time consuming, so it is challenging to learn a model for the target information. Whereas we have a large amount of unlabeled data which could be used to learn a model for target domain. In that respect are various existing approaches for face recognition that uses few labeled and a heavy amount of unlabeled data for facial expression identification. In that respect are several approaches is that update the theoretical account with fewer labeled and large unlabeled data, but these algorithms take the whole unlabeled data in up front.

In our framework we perform the transfer learning in an online manner with the sequential incremental unlabeled data available from the target domain. In our framework transfer learning is enabled using co training; we are using online co-training approach by Bhatt et al. to update the classifiers decision boundary. To the best of our knowledge the first algorithm that uses transfer learning for face recognition cross matching as a semi supervised method. This generalized framework could be applied to any classifier that allows the retraining with the available incremental data. We apply our concept of co training to SVM classifiers, where the SVM classifier is trained using the available initial training data and the decision hyper plane for the classifier is obtained. Then the SVM classifiers are being updated using the available data and the previous support vectors of the classifier.

In face recognition, the classifiers such as SVM, are learned using training data (from the source domain) while the performance is evaluated on a separate unseen test data (the target domain) which may have different behavior and properties and follow a different distribution compared to the training data. Consider a scenario where there are two classifiers, one trained using the source and other trained using the target domain data. During training, there is a large amount of labeled data in the source domain, i.e., for matching HR probe with HR gallery images (source domain) but only a few labeled instances are available in the target domain, i.e., for matching LR probes with HR gallery images. In such a scenario, the source domain classifier alone may not, efficiently classify the test instances because of the variations in data distrib

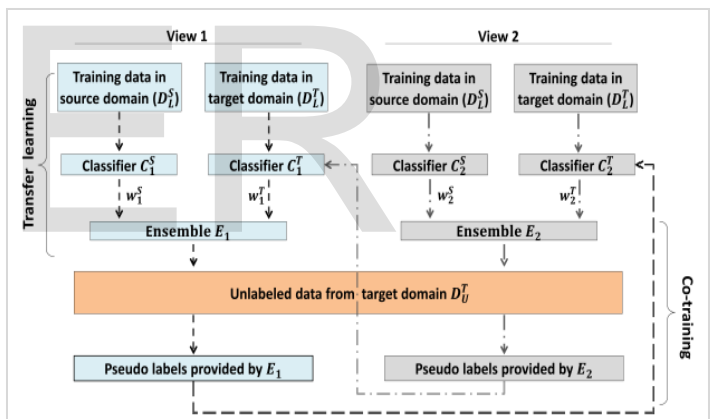


Figure 5: Block diagram for ensemble based Co transfer learning framework

ution of source and target domains. Since the classifier in target domain is trained using only a few labeled samples, it is not able to efficiently classify the test instances. It has to learn/update its decision boundary with the incremental data available in the target domain. Both the classifiers are individually insufficient to classify the test data from the target domain. Consequently, in the proposed algorithm, an ensemble is constructed as a weighted combination of the source and target domain classifiers. It efficiently classifies test instances and subsequently transfers the knowledge from the source domain to the target domain as and when the data from the target domain is available. For this, the two classifiers trained on the source and target domains are combined to efficiently classify the unlabeled probe instances.

As recorded in Figure 5, the source domain classifiers (C_j^s) are trained using sufficient HR labeled training data denoted by (D_L^S)={(u_1^s,z1), (u_2^s, z2),……. ($u_{n,}^s z_n$)}.Every ith instance u_i has two views $x_{i,1}$,$x_{i,2}$ for the training label zi € {-1,+1};here xi,1

and xi,2 represent the input vectors obtained from two separate features.{-1} refers to the impostor class where the query and probe images belong to different subjects and{+1} refers to the genuine class where the gallery and the probe images belong to the same subject. The two features are utilized for co-training target domain classifiers C_j^T are initially trained on a few labeled instances from the target domain represented as (D_L^T)={$(u_1^T, z1)$, $(u_2^T, z2)$,……. $(u_{m,}^T z_m)$}.. Here n, m are the no of training instances in the source and target domain respectively such that n>m and j=1,2 represents the feature .Let a set of r unlabeled probe instances in the target domain can be represented as (D_U^T)=u'^T_1,An ensemble prediction denoted as EJ is constructed for each view. Ej is a weighted combination of the source domain and target domain classifiers with wi, jS and wi, jt are the weights of the source domain and target domain classifier for the ith and jth view. For the ith unlabeled probe instance in the jth view, the ensemble function predicts the label. EJ (xi,j)->yi,j. For the ith instance in the target domain class label is predicted by the ensemble as given in Equation 1

$$y_{i,j}=sign(w_{ij}^S \pi\left(C_j^S(u'_i)\right) + w_{i,j}^T \pi\left(C_j^T(u'_i)\right) - \frac{1}{2}$$

Where π is a normalization function such that π(x)=max(0,min(1,x+1/2)) and initially the weights for the source and target domain classifiers at the ith instance is set to 0.5,so that each classifier contributes equally within an ensemble. Gradually these are automatically adjusted to emphasize the contribution from updating target domain classifiers in an ensemble. As proposed by Zhao and Hoi, the weights are updated dynamically

Co-training: As we discussed, we have an abundance amount

$$w_{i+1,j}^S = \frac{w_{i,j}^S h_i(C^S)}{w_{i,j}^S h_i(C^S) + w_{i,j}^T h_i(C_j^T)}$$

$$w_{i+1,j}^T = \frac{w_{i,j}^T h_i(C^T)}{w_{i,j}^S h_i(C^S) + w_{i,j}^T h_i(C^T)}$$

of unlabeled probe instances which could be used to update or to learn classifiers in the target domain. Obtaining labeled training instances of the target domain is expensive and a difficult task. There are situations in biometrics when there is just a modest quantity of labeled data is usable for training while an ample amount of unlabeled data is useable as a probe. In such a post co-training is proven to be beneficial as it can be used to assign pseudo labels to large unlabeled instances. In the proposed co training, it assumes two ensemble classifiers E1 and E2 are trained on two separate views where each ensemble classifier has sufficient accuracy of prediction. If the first ensemble confidently predicts candid label for an instance while other predicts impostor label with a low confidence of prediction, then this instance is being used for updating the second ensemble and if the second ensemble confidently predicts candid label for an instance and first ensemble predicts an impostor label with a low confidence of prediction, then this particular instance would be used to retrain the first ensemble function. The confidence of prediction for an instance on the jth view is denoted by αj, which is measured as distance of the instance from the decision hyperplane. The confidence of prediction required for an instance to belong to the candid class, if the distance from decision boundary should be greater than the candid threshold (PJ). Likewise, an instance is confidently predicted as impostor class, if the distance from decision boundary is greater than the impostor threshold (Pj).
Since we are using SVM classifier a candid threshold is computed as the distance of the farthest support vector of candid class. An impostor threshold is computed as the distance of farthest support vector of impostor class. By varying the thresholds, will change the no of instances used for co-training .Higher values of threshold give conservative co training and smaller value leads to aggressive co training .By this way the large amount of unlabeled instances is transformed into pseudo labeled instances, which are being used for updating the ensemble function of the classifiers. The main objective of selecting two ensembles is to enable Co transfer learning as one ensemble function provides pseudo labeled training instances to other.

Input: Initial labeled training data D_L^S in the source domain, a few labeled instances D_L^T from the target domain. Unlabeled probe instances D_U^T from target domain (available sequentially).
Iterate: j= 1 **to** 2 (number of views)
Process: Train classifiers C_j^S and C_j^T on j^{th} view of D_L^S and D_L^T respectively to construct ensemble E_j. Compute confidence thresholds P_j for each view.
for i = 1 **to** r (number of probe instances) **do**
  Predict labels: $E_j(x_{i,j}) \rightarrow y_{i,j}$; calculate α_j: confidence of prediction
  if $\alpha_1 > P_1$ & $\alpha_2 < P_2$ **then**
    Update C_2^T with pseudo-labeled instance $\{x_{i,2}, y_{i,1}\}$ & recompute w_2^S and w_2^T.
  end if.
  if $\alpha_1 < P_1$ & $\alpha_2 > P_2$ **then**
    Update C_1^T with pseudo-labeled instance $\{x_{i,1}, y_{i,2}\}$ & recompute w_1^S and w_1^T.
  end if.
end for.
end iterate.
Output: Updated classifiers C_1^T, C_2^T and weights w_1^S, w_1^T, w_2^S and w_2^T.

Figure 5: Algorithm for co-transfer learning

Co-transfer: In the framework, transfer learning and co training are used in a non-separable way to improve the target domain face matching with pseudo labels being assigned by co training, which leads to transfer of knowledge learnt from the source domain to target domain. Within each ensemble function, the target domain classifier updates its decision boundary, with the pseudo labeled obtained during the testing phase.

The corresponding weights of source domain and target domain are adjusted dynamically using the equations given below. By dynamically updating the weights avoids the need to determine the target domain classifiers from the start and get the advantages of system scalability and computational efficiency. The target domain classifiers are only updated with the pseudo labeled instances, whereas the source domain does not update since they are being easily prepared with an ample quantity of labeled data which is being available up front in the source domain[1]. The algorithm got co-transfer learning framework is being given in the algorithm.

To evaluate the effectiveness of the co –transfer learning algorithm, we find the error bounds for each ensemble. For an ensemble E, M_E errors of the ensemble are bounded by

For two ensembles function the final decision of classification is based on their combination,

$Min(M_{E1},M_{E2}) <= M <= Max(M_{E1},M_{E2})$……(3)

where M is the error bound for co-transfer learning framework. Co-transfer learning saturate, as more and more pseudo labeled instances are available, and then importance is being shifted to target domain classifiers[6].

4. CO-TRANSFER LEARNING FRAMEWORK FOR CROSS RESOLUTION FACE MATCHING.

In an operational scenario, training is performed in a controlled environment, during testing phase it encounters data from uncontrolled situations which has different data distributions and properties. For recognizing cross resolution face images, co-training is particularly useful. The source domain and the target domain classifiers are trained on two different views. One view is local binary pattern (LBP) and the other view is the combination of histogram of oriented gradients (HOG) and scale invariant feature transform (SIFT). These views are being resilient to scale changes and variations in illumination. These two features provide the information which is diverse.

Local Binary Pattern (LBP): It is one of the best feature descriptor for face recognition, since face is a composition of micro patterns[9]. These micro patterns can be well described using the LBP operator. The operator works by assigning a label to each pixel of an image, by thresholding the 3*3 neighborhood of each pixel with the center pixel values. By concatenating the eight cells to 8 bit code gives the code for the center pixel. Later the LBP operators were extended to use neighborhood of different images. Thus a circle was made with radius R from the center pixel and the neighborhood size of P equal space pixels of the circle, thus it gives the LBP operator LBP (P, R). LBP is invariant to monotonic photometric change and can be efficiently extracted[5].

Scale invariant Feature Transform (SIFT): It is a scale and rotation invariant descriptor that gives a compact representation of an image based on magnitude, orientation and spatial vicinity of image gradients[4]. SIFT proposed by Lowe is a sparse descriptor and it can be used equally in a dense manner also. SIFT descriptors are computed at the sampled regions which are then concatenated and chi square distance is used to compare two SIFT descriptors.

H*istogram of Oriented Gradients (HOG):* Are feature descriptors for the purpose of object detection. It counts the occurrences of gradient orientation in localized portions of the image. By HOG descriptors, the local object appearance and shape of an image can be traced by a distribution of intensity gradients or by means of edge directions. For improved accuracy, the local histograms can be contrasted normalized. By normalizing it gives better invariance to changes in illumination or shadowing.

The use of low-level feature descriptors has been an effective approach in face recognition. The scale - invariant feature transform (SIFT) and histogram of oriented gradients (HOG), which can be viewed as a quantized code of the facial gradients, is used in face recognition as effective descriptors. By blending the results of SIFT and Hog gives better result and helps to cut the feature length.

4.1 Initial training on labeled data from source and target domains

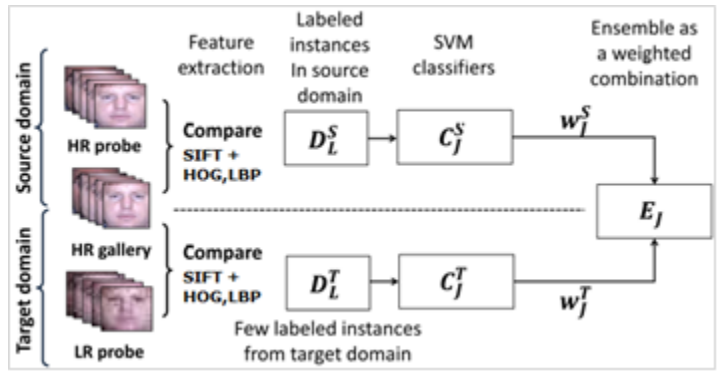


Figure 6: Training process of source and target domain classifiers to build an ensemble

The co-transfer learning framework assumes that during training, each field has a high resolution gallery-probe pairs and a few subjects have corresponding low resolution images from the target area. Face images are tessellated into non-overlapping facial patches[7]. LBP and SIFT and HOG descriptors are computed for each local patch and matched using the distance measure. Distance scores corresponding to each local patch are vectorized to an input vector, where is the associated label. +1 signifies that the gallery-probe pair belongs to the same individual (i.e. genuine pair) whereas -1 signifies that the gallery- probe pair belongs to images corresponding to different individuals (i.e. impostor pair).Input vectors obtained by matching LBP descriptors of two high resolution images are utilized for training the source domain Support Vector Machine (SVM) classifier on view1.On the contrary, the target domain Support Vector Machine (SVM) classifiers for view 1 are trained using one high resolution and one low resolution images. The source domain and the target domain Support Vector Machine (SVM) classifiers are then combined to form an ensemble. Similarly, the Support Vector Machine (SVM) classifiers for view 2 (SIFT) are trained and the ensemble function is learned.

4.2 Co-transfer learning with unlabeled probes from target domains

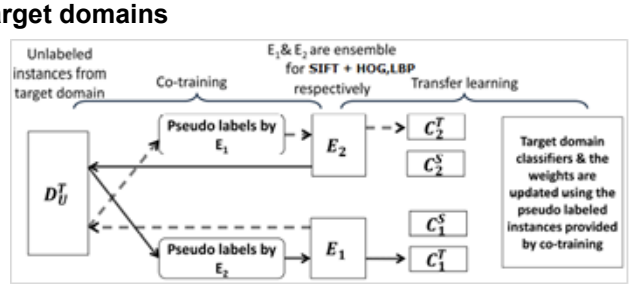


Figure 7: Co-transfer in target domain with unlabeled instanes.

For matching a Low Resolution probe with a High Resolution Gallery image, the images are tessellated into non-overlapping local patches and LBP, HOG and SIFT descriptors is computed for each local patch. LBP descriptors from the corresponding local patches on the gallery and probe images are matched using distance and the distance scores from these local patches are vectorized to form an input vector u for view 1. Similarly, an input vector corresponding to SIFT and HOG (view 2) is computed using the distance measure. Unlike training, the instances obtained during testing are unlabeled. For every query given to the biometric system, both the ensembles and are used to classify the instance. If one ensemble confidently predicts impostor label with low confidence while the other ensemble predicts genuine label for an instance, then this instance is added as a labeled re-training sample for the second ensemble and vice-versa. The target domain Support Vector Machine (SVM) classifiers in the ensembles are updated with pseudo-labeled probe instances obtained during testing. Further, the weights for both source domain and target domain Support Vector Machine (SVM) classifiers are also updated with each pseudo-labeled probe instance. Thus, each ensemble updates the target domain classifier of the other ensemble. The final decision is computed by combining responses from both the ensembles.

5. EXPERIMENTAL RESULTS

5.1 Database

The performance of the proposed CTL is evaluated on three different bases 1)SCface 2)CMU MultiPIE 3)Choke Point. The experiments are being planned in order to contemplate the real world scenarios where there is an ample sum of training ample data for training the source domain. However, only a few low resolution probe images and its corresponding high resolution images, for training the target domain classifiers. The training subjects in target domain are a subset of the training subjects in the source domain. Details about the database are further described 1) SCface: It is a surveillance database comprising images of 50 individuals captured in an uncontrolled environment, using multiple surveillance placed at different locations and distance. 2) CMU Multipie: it consists of 75 individuals captured in four different ways with varying pose, expression and illumination. The individuals with frontal pose and neutral expression are being selected. For each subject, one high resolution image is kept in the gallery and one low resolution is used as a probe. 3) Choke Point database: Images are captured by surveillance cameras in uncontrolled environments and include illumination, expression and pose variations. The database consists of 5 unique subjects with five variations for each subject.

To equate the conditions that the gallery is generally captured under controlled conditions, the resolution of gallery pictures is constantly higher than the probe images. Experiments are done at different resolutions of the gallery and probe images ranging from 216*216 pixels to 32*32 pixels.

5.2 Results and Analysis

For cross resolution face matching, the positive potential of algorithms degrade because of divergence in information content between high resolution gallery and low resolution probes, limited biometric information in low resolution probe images. The framework addresses these issues by using the knowledge learned during the training phase. The main goal of experiments to conclude the effectiveness of the proposed algorithm in transferring knowledge from the source domain to target domain. It also validates our assertion that co-training enables updating the decision boundary of the target domain classifiers with unlabeled probe instances as and when they arrive. 1) Cross-pollination of transfer learning and co-training seamlessly transfers the knowledge discovered in the source domain for matching cross-resolution face images. Co training and transfer learning go hand-in-hand as co-training provides pseudo labels for unlabeled test instances which in-turn are used to update the target domain classifiers within each ensemble and thus transfer the knowledge. 2) Updating the weights of the source and target domain classifiers allows dynamically adapting the donation from the constituent source and targeting domain classifiers in an ensemble. Initially, equal weights are assigned to both the classifiers; however, with knowledge transfer, weights of classifiers in the target domain become more salient.

Co-training provides correct pseudo labels for about 96 percent of the total instances. The poor performance can be attributed to the fact that some of the pseudo labels assigned to unlabeled probe instances may be incorrect, leading to negative transfer. However, the effect of negative-transfer can be minimized by optimally selecting the confidence threshold for co-training. High threshold value implies conservative transfer while a smaller value of the threshold contributes to aggressive transfer.

5.3 Performance Evaluation of SVM classifier

Efficiency or accuracy of the SVM classifier is evaluated based on the error rate. This error rate can be delineated by the terms true positive and false positive and true and false negative as follows:

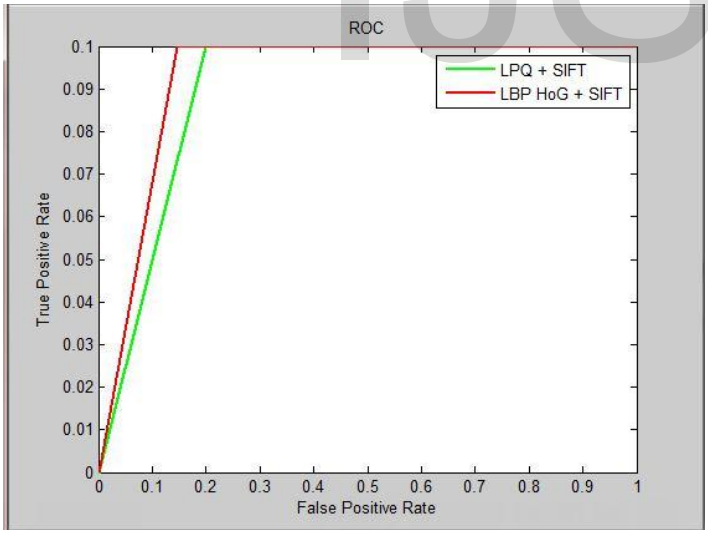
- True Positive (TP): Correctly identified.
- False Positive (FP): Incorrectly identified.
- True Negative (TN): Correctly Rejected.
- False Negative (FN) : Incorrectly Rejected.

Sensitivity (True Positive Rate) measures the ratio of positives that are correctly distinguished as such (e.g., the percentage of sick people who are correctly distinguished as accepting the condition). Specificity (True Negative Rate) measures the proportion of negatives that are correctly distinguished as such (e.g., the percentage of healthy people who are correctly distinguished as not accepting the condition).

Applying the above concepts, based on the training and test images and also based on the probe instances assigned, we have estimated the class operation which is being shown be-

low. The outcome is a comparison with the existing scheme which utilizes the feature extractors (LPQ AND SIFT) and with our proposed scheme which utilizes the feature extractors (LBP AND SIFT+HOG).

For any test, there is usually a trade-off between the measures. This tradeoff can be represented graphically as a receiver operating characteristic curve. A receiver operating Characteristic (ROC), or ROC curve, is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is made by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. The true-positive rate is also known as sensitivity, or recall in machine learning. The false-positive rate is also known as the fall-out and can be calculated as (1 - specificity). The ROC curve is thus the sensitivity as a function of fall-out. ROC analysis provides tools to select possibly optimal models and to discard suboptimal ones independently from (and prior to specifying) the cost context or the class distribution. ROC analysis is related in a direct and natural way to cost/benefit analysis of diagnostic decision making. The Area Under ROC curve (AUC) has been used to determine the over-all classification accuracy. By calculating AUC, we can measure the class discrimination capability of a specific classifier. An area of above 0.5 represents a perfect test while an area of less than or equal to 0.5 represents worth less test. The larger the area (the higher AUC value) means higher the classification performance. In this project, the ROC analysis and accuracy are used to measure the performance of the classifiers.



6. Conclusion

Face recognition is a both challenging and important recognition technique. Among all the biometric techniques, face recognition approach possesses one great advantage, which is its user-friendliness (or non-intrusiveness). However, there are several challenges that are involved in creating an efficient facial recognition system, such as performance in low light, different resolutions, etc. The project has tried to focus the efforts on issues caused by cross resolution images. Co-transfer learning framework which seamlessly combines the co-training and transfer learning paradigms for effective cross-resolution face matching. During training, the framework learns to match high resolution face images in the source domain. This cognition is then transmitted from the source area to the target area to match low resolution probes with high resolution gallery. The framework builds ensembles from the weighted combination of source and target domain classifiers on two separate views. Two ensembles trained on separate views transform the unlabeled probe instances into pseudo-labeled instances using co-training. These pseudo labeled instances are utilized for updating the decision boundary of the target domain classifier, thus, transferring knowledge from the source domain to the target domain. Further, dynamically updating the weights assigned to each classifier facilitates the gradual shift of knowledge from the source to target domain. The amalgamation of transfer learning and co-training helps to transfer the knowledge from the source to target domain with probe instances as and when they arrive. The Co-transfer learning framework provides significant improvement in cross-resolution face matching on different surveillance quality face databases.

References

[1] Himanshu S. Bhatt, Richa Singh, Mayank Vatsa, and Nalini K. Ratha, `Improving Cross-resolution Face Matching using Ensemble based Co-Transfer Learning', IEEE Transactions on Image Processing, Volume 23, Issue 12, pp. 5564-5669, 2015.

[2] H. Hua ng and H. He, `Super-resolution method for face recognition using nonlinear mappings on coherent features', IEEE Transactions on Neural Networks, vol. 22, no.1, pp. 121-130, 2011.

[3] K. Jia and S. Gong, `Multi-modal tensor face for simultaneous superresolution and recognition', in Proceedings of International Conference on Computer Vision, 2005, pp.1683-1690.

[4] David G. Lowe, `Distinctive image features from scale-invariant keypoints.', International journal of computer vision, pp 145-200,2004.

[5] V. Ojansivu and J. Heikkila. , `Blur insensitive texture classification using local binary patterns' In ICISP ,pp 150-180, 2008.

[6] H. S. Bhatt, S. Bharadwaj, R. Singh, M. Vatsa, A. Ross, and A. Noore, `On co-training online biometric classifier', in Proceedings of Inter- national Joint Conference on Biometrics, 2011, pp 1-6.

[7] P. H. Hennings-Yeomans, S. Baker, and V. Bhagavatula, `Simultaneous super-resolution and feature extraction for recognition of low-resolution faces', in Proceedings of Computer Vision and Pattern Recognition, 2008, pp. 1-8.

[8] W. W. W. Zou and P. C. Yuen, `Very low resolution face recognition problem', IEEE Transactions on Image Processing, vol. 21, no. 1, pp. 327-340, 2012.

[9] Z. Lei, T. Ahonen, M. Pietikainen, and S. Li, `Local frequency descriptor and local binary pattern for low-resolution face recognition', in Proceedings of International Conference on Automatic Face Gesture Recognition and Workshops, 2011, pp. 161-166.

[10] B. Li, H. Chang, S. Shan, and X. Chen, `Low-resolution face recognition via

IJSER

coupled locality preserving mappings', IEEE Signal Processing Letters, vol. 17, no. 1, pp. 20-23,2010.

IJSER